

融合注意力机制与全局信息的 SAR 舰船目标检测算法

赵晶钰, 李 敏, 陈谢发, 吕奕龙, 何玉杰
(火箭军工程大学, 陕西 西安 710025)

摘 要: 为提高合成孔径雷达图像中舰船目标的检测精度, 基于 YOLOv8 算法框架提出了一种融合感受野注意力机制与全局信息的目标检测算法 RA-YOLO (Receptive-field Attention YOLO)。为了增强对重要信息的感知能力, 在骨干网络中引入感受野注意力卷积; 同时, 在空间金字塔池化结构中加入全局最大池化层和全局平均池化层, 融入背景信息和边缘信息以提高检测精度; 最后, 引入内部交并比损失函数提升网络的泛化能力, 以提高回归速度和预测框的准确度。实验结果表明: 在公开数据集上与 YOLOv8 算法相比, 在参数量和运算速度几乎不变的情况下, 所设计算法的平均精度提高了 3.0 个百分点; 综合性能优于其他经典目标检测算法。

关键词: 舰船目标检测; 深度学习; YOLOv8 算法; 合成孔径雷达; 感受野注意力

中图分类号: TP391.4

文献标志码: A

文章编号: 2097-2741(2024)04-040-007

SAR Ship Target Detection Algorithm Combining Attention Mechanism and Global Information

ZHAO Jingyu, LI Min, Chen Xiefa, LYU Yilong, HE Yujie
(Rocket Force University of Engineering, Xi'an 710025, Shaanxi)

ABSTRACT: In order to improve the detection accuracy of ship targets in synthetic aperture radar (SAR) images, a target detection algorithm RA-YOLO integrating the receptive field attention mechanism and global information was proposed based on YOLOv8 algorithm framework. In order to enhance the important information perception ability, the Receptive-Field Attention convolution was introduced into the backbone network. At the same time, a global maximum pooling layer and a global average pooling layer were added to the spatial pyramid pooling structure to integrate background information and edge information, by which the detection accuracy was improved. Finally, the Inner-IoU loss function was introduced to improve the generalization ability of the network. Thereby, the regression speed and the accuracy of the prediction frame were improved. Results of experiments on open data set showed that when the number of parameters and operation speed are almost unchanged, the average accuracy of the proposed algorithm is 3% higher than that of YOLOv8 algorithm, and the comprehensive performance is better than other classical target detection algorithms.

收稿日期: 2024-04-27

第一作者: 赵晶钰 (1996—), 男, 硕士研究生, 主要从事图像生成与目标识别研究。E-mail: 369360770@qq.com

引用格式: 赵晶钰, 李敏, 陈谢发, 等. 融合注意力机制与全局信息的 SAR 舰船目标检测算法[J]. 火箭军工程大学学报, 2024, 38(4): 40-46.

KEYWORDS: ship target detection; deep learning; YOLOv8 algorithm; synthetic aperture radar; receptive-field attention

随着海洋权益逐渐成为大国之间博弈的焦点,对海域中的舰船进行监视预警和目标锁定对维护国家海洋安全十分重要。合成孔径雷达(Synthetic Aperture Radar, SAR)成像是一种无损探测技术,它可以不受光照、云层、气候等多种条件的影响,提供高分辨率的图像^[1],是获取海域图像信息的重要渠道之一,基于深度学习的SAR图像目标检测已经成为当前的研究热点^[2]。在军事领域中,SAR图像舰船目标检测被应用到敌舰侦察、精确制导等诸多领域。然而,受限于SAR图像的固有噪声以及舰船目标范围较小、类内尺度不一等问题,直接使用经典目标检测算法进行检测往往效果十分有限。

目标检测是计算机视觉领域的核心研究方向,基于深度学习的目标检测算法大致可分为两阶段目标检测算法与一阶段目标检测算法。两阶段目标检测算法主要以基于区域的卷积神经网络(Region-based Convolutional Neural Networks, RCNN)^[3]、Faster RCNN^[4]、特征金字塔网络(Feature Pyramid Networks, FPN)^[5]等为代表,这类目标检测算法精度高、效果好,但速度慢,实时性低。一阶段目标检测算法以单激发多框探测器(Single Shot Multibox Detector, SSD)^[6]、你只看一次(You Only Look Once, YOLO)^[7]为代表,这类算法计算量较小,实时性高。其中,YOLO系列算法因其优秀的性能被广泛应用到各类目标检测任务中,但受限于目标特性不同,效果参差不齐。

目前,国内外许多学者基于YOLO系列算法对SAR图像舰船目标检测进行了一系列研究。张上等^[8]在YOLOv7算法的基础上,从网络结构、锚框参数、注意力机制等方面进行了改进,提升了对舰船目标的检测效率。Zhao等^[9]提出了一种新的SAR舰船检测算法。通过在骨干网络中嵌入Swin Transformer网络架构,并利用带有残差结构的路径聚合网络构建特征金字塔,提高了模型的收敛速度和检测精度。Guo等^[10]提出了一种基于YOLO的轻型多尺度舰船探测器,使用改进的S-Mobilenet模块作为主干特征网络,在不增加额外计算的情况下增强了特征提取能力,并通过改进损失函数等方法,提高了模型的检测精度。

上述方法均可以在一定程度上解决YOLO算法在SAR舰船目标检测上的性能缺陷,但仍然存在一些不足。考虑到舰载、弹载设备的算力有限,且对算法实时性要求较高,本文在前人研究的基础上,基于YOLOv8算法框架,提出了一种融合注意力机制与全局信息的SAR舰船目标检测算法RA-YOLO(Receptive-Field Attention YOLO),可在保证实时性的前提下提升对舰船目标的检测能力。

1 SAR舰船目标检测算法RA-YOLO

1.1 整体框架

YOLOv8采用了轻量级网络结构,具有出色的检测性能,它的基本框架主要包括主干网络(Backbone)、颈部网络(Neck)和头部模块(Head)。针对SAR图像噪声明显、分辨率相对较低,且舰船目标普遍较小的特点,本文以YOLOv8为基线模型,对算法网络架构进行了优化设计,从卷积、空间金字塔池化、损失函数等方面对算法进行改进,提出了一种SAR图像舰船目标检测算法RA-YOLO,其网络结构如图1所示。

1.2 感受野注意力卷积

注意力机制(Attention Mechanism)在目前计算机科学和深度学习领域中广为流行^[11],它可以使算法在对数据进行处理时关注到更多的重点内容。感受野注意力(Receptive-Field Attention, RFA)^[12]是一种注意力机制,它的特点是能够更加充分地考虑不同特征的重要程度,有效地解决目前空间注意力机制中卷积核参数共享的困境。基于上述感受野注意力机制思想设计的感受野注意力卷积——RFACnv,它强调了接收域滑块中不同特征的重要程度,以此对接收域的空间特征进行优先级排序,从而有效处理了常规卷积对位置变动感知不敏锐的问题。

RFACnv将接收的信息首先通过平均池化进行处理,然后使用 1×1 卷积运算进行信息交互,将所得结果与经过标准 3×3 卷积计算后得到的结果进行重组。相较于原始算法,RFACnv小幅增加了参数量和计算量,但显著提升了检测精度。其结构如图2所示。

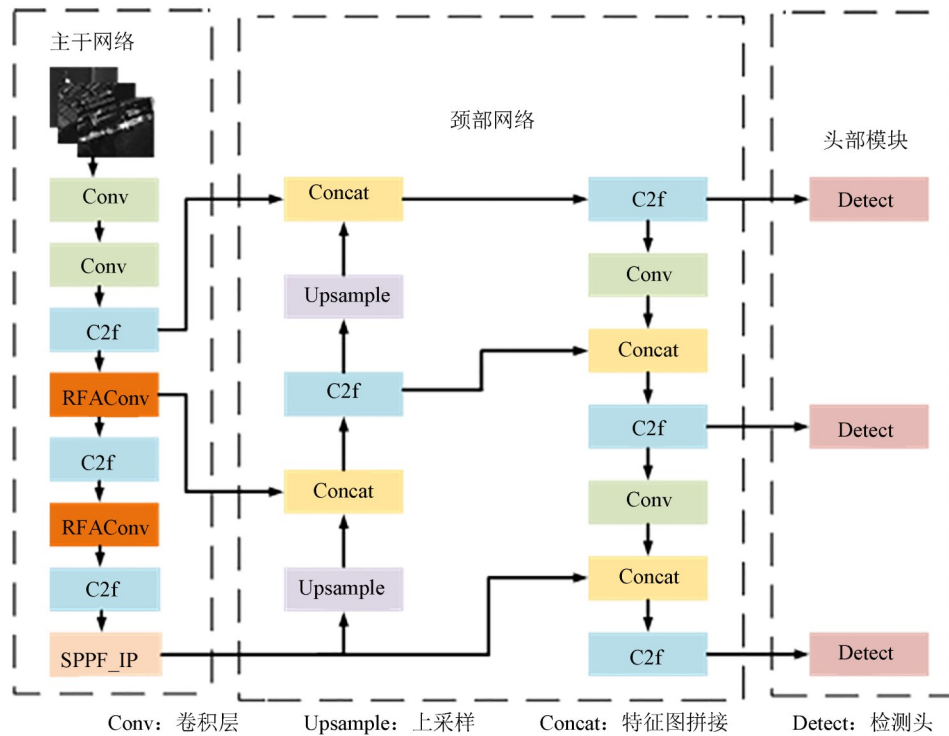


图 1 RA-YOLO算法结构

Fig. 1 Structure of RA-YOLO algorithm

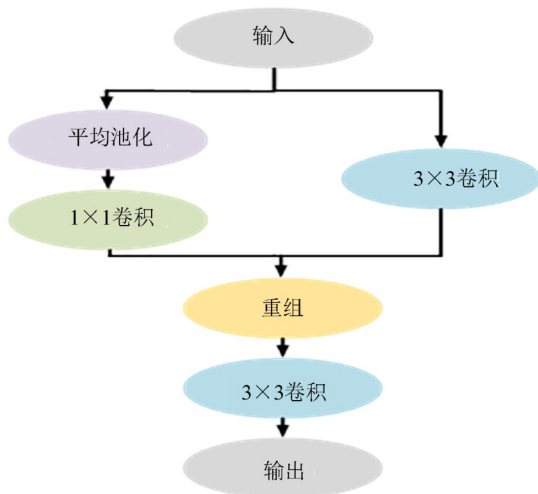


图 2 RFACnv结构

Fig. 2 Structure of RFACnv

本文将YOLOv8骨干网络中的两层Conv改进为RFACnv，增强了骨干网络在进行特征提取时对重要因素的感知能力，优化了网络的特征提取能力，提升了算法的检测性能。

1.3 融合全局信息的特征金字塔模块

在SAR图像舰船目标检测任务中，图像中只含有海洋背景和舰船目标，因此，准确处理边缘和背景信息的关注度有利于算法性能提升。YOLOv8算法通过使用空间金字塔池化结构（Spatial

Pyramid Pooling Fast, SPPF）使输入数据依次经过3个最大池化层，然后将输出拼接起来得到多尺度特征，实现全局特征与局部特征的融合。

考虑到SPPF主要实现了对边缘信息的关注，而对背景信息缺乏必要的重视。本文对SPPF进行改进重构，并联了一个全局最大池化层和一个全局平均池化层，然后将结果进行拼接，形成了一个模块SPPF-IP，以增强全局视角信息，帮助网络更好地做出判断。SPPF和SPPF-IP的结构如图3所示。

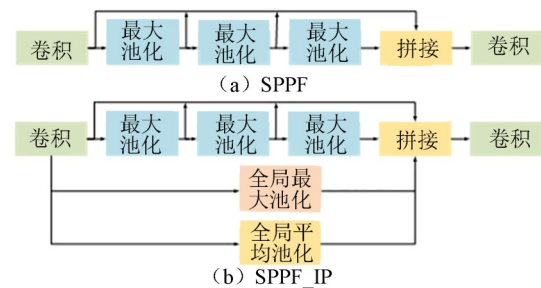


图 3 SPPF和SPPF_IP模块的结构

Fig. 3 Constructions of SPPF and SPPF_IP

1.4 内部交并比损失函数

损失函数在目标检测算法中的应用尤其关键，YOLOv8通过使用CIoU^[13]（Complete Intersection over Union）损失函数来计算预测框的损失。CIoU

向交并比 (Intersection over Union, IoU) 损失函数添加了几何约束, 以加速收敛和提升检测性能。但它并未考虑 IoU 损失函数本身的合理性, 在一定程度上影响了检测结果的质量。

通过区分不同的回归样本, 并使用不同尺度的辅助边界框来计算损失, 可以有效加速边界框回归过程。因此, 本文引入了 Inner-IoU^[14] 损失函数, 它使用辅助边界框来加速回归, 并通过尺度因子来控制辅助边界框的尺度大小, 从而克服了现有损失函数在泛化能力方面的局限, 提高了预测框的准确度。其计算公式如下:

$$\begin{aligned}
 b_{(gt,l)} &= x_{(gt,c)} - \frac{w_{gt} r}{2} \\
 b_{(gt,r)} &= x_{(gt,c)} + \frac{w_{gt} r}{2} \\
 b_{(gt,t)} &= y_{(gt,c)} - \frac{h_{gt} r}{2} \\
 b_{(gt,b)} &= y_{(gt,c)} + \frac{h_{gt} r}{2} \\
 b_l &= x_c - \frac{wr}{2} \\
 b_r &= x_{(gt,c)} - \frac{wr}{2} \\
 b_t &= y_c - \frac{hr}{2} \\
 b_b &= y_{(gt,c)} - \frac{hr}{2} \\
 S_{inter} &= (\min(b_{(gt,r)}, b_r) - \max(b_{(gt,l)}, b_l)) \cdot \\
 &\quad (\min(b_{(gt,b)}, b_b) - \max(b_{(gt,t)}, b_t)) \\
 S_{union} &= (w_{gt} h_{gt}) (r)^2 + \\
 &\quad (wh) (r)^2 - S_{inter} \\
 R_{inner-IoU} &= \frac{S_{inter}}{S_{union}}
 \end{aligned}$$

式中: b 和 b_{gt} 分别表示锚框和目标框; x_c 和 y_c 为锚框的中心点; $x_{(gt,c)}$ 和 $y_{(gt,c)}$ 为目标框的中心点; w 和 h 为锚框的宽度和高度; w_{gt} 和 h_{gt} 为目标框的宽度和高度; b_l 、 b_r 、 b_t 、 b_b 是使用尺度因子生成的辅助边框的坐标; r 为可调节的比例因子。

2 实验结果与分析

2.1 实验环境

实验使用的硬件平台配置为: 处理器 Inter core i7-14 700 K, 显卡 NVIDIA TITAN RTX, 显存 24 GB, 内存 64 GB。使用 Anaconda 环境、python3.9 和 cuda11.8。训练设置采用随机梯度下降法 (Stochastic Gradient Descent, SGD), 初始

学习率为 0.01, 输入图像大小为 320 像素 $\times$ 320 像素, batchsize 设置为 2, 迭代次数为 50 次, 早停轮数设置为 10。

2.2 数据集介绍

采用海军航空航天大学提出的公开 SAR 舰船数据集 SSDD (SAR Ship Detection Dataset)^[15] 进行实验。数据集中共有 1 160 张图像, 其中包含不同分辨率、尺寸、海况下采集的舰船 SAR 图像, 平均每个图像有 2.12 个舰船目标。实验中将数据依照 7:2:1 的比例划分为训练集、验证集和测试集进行实验。

2.3 评价指标

使用目标检测领域常见的准确率 P 、召回率 R 、平均精度 P_{mAP50} (IoU 阈值为 0.5 时的平均准确度) 作为模型性能评价指标, 计算公式为

$$\begin{aligned}
 P &= \frac{\eta_{TP}}{\eta_{TP} + \eta_{FP}} \\
 R &= \frac{\eta_{TP}}{\eta_{TP} + \eta_{FN}} \\
 P_A &= \int_0^1 P(R) \\
 P_{mAP} &= \frac{1}{k} \sum_{i=1}^k P_{A,i}
 \end{aligned}$$

式中: η_{TP} 、 η_{FP} 、 η_{FN} 分别表示正确检测、错检测、漏检测目标; P_A 是识别正确的样本数占识别样本总数的百分数, 对应着 $P(R)$ 曲线下方的面积; P_{mAP} 是所有类别 P_A 值的平均数, 通常将 P_{mAP} 作为评估模型性能的最终指标。

此外, 为了证明算法的实时性, 还引入了每秒 10 亿次浮点运算数、参数量作为评价指标。

2.4 消融实验

在 SSDD 数据集上对相关改进模块进行了消融实验, 实验结果如表 1 所示。

通过表 1 可以看出, 每种模块的加入都不同程度地提高了算法的性能。在加入 RFACnv 后, R 和 P_{mAP50} 分别提高了 5.2% 和 1.7%, 说明 RFACnv 能够充分考虑感受野中不同特征的重要性, 使注意力聚焦于感兴趣的区域。加入 SPPF_IP 后, R 和 P_{mAP50} 分别提高了 5.2% 和 1.7%, 证明强调对背景的关注后, 有利于提高算法性能。使用 Inner-IoU 损失函数后, R 提高了 6.9%, P_{mAP50} 提高了 2.1%, 说明该损失函数有效提高了模型的检测精度, 且不会增加运算成本。将所有模块结合后, 算法 P 维持在原有水平, R 和 P_{mAP50} 分别取得了 8.2% 和 3% 的提升, 参数量和计算复杂度略微增

加，但不影响模型的实时性。

2.5 检测效果可视化

将 YOLOv8 和本文改进算法 RA-YOLO 在相同实验条件和参数下进行训练，然后选取一些 SAR 舰船图像进行检测，检测效果如图 4 所示。

可以看出：在对离岸和近岸稀疏舰船目标进行检测时，虽然基线算法和本文改进算法都成功检测出了目标，但本文算法的置信度更高；在对密集舰船目标进行检测时，改进算法检测出了 8 只舰船，而基线算法只检测出了 4 只舰船，且改进算法的置信度更高。这充分证明了本文提出的改进算法的检测能力。

2.6 对比实验

选取目标检测领域中有代表性的算法 SSD、

RT-DETR^[16](Real-time Detection Transformer)、YOLOv5、YOLOv6、YOLOv7 进行对比实验。其中，YOLOv5 和 YOLOv7 选择了和本文算法计算规模相近的 YOLOv5s 和 YOLOv7tiny 进行比较，具体检测结果如表 2 所示。

可以发现，本文改进算法相较于其他算法具有更好的检测性能。相较于 SSD 和 YOLO 各系列算法，本文提出的算法在较少的参数量和运算数的情况下具有较高的准确率。RT-DETR 虽然在 P_{mAP50} 上略微领先本文所提出的算法，但参数量和浮点运算数均数倍提升，综合性能落后于本文提出的改进算法。在 SAR 图像舰船目标检测场景中，本文提出的算法展现出了较为明显的优势，证明了其有效性和先进性较好。

表 1 消融实验

Tab. 1 Results of ablation experiments

RFACnv	SPPF_IP	Inner - IoU	P/%	R/%	$P_{mAP50}/\%$	参数量/个	每秒 10 亿次浮点运算数/次
\	\	\	0.898	0.712	0.824	3 005 843	8.1
√	\	\	0.868	0.764	0.841	3 026 963	8.1
\	√	\	0.879	0.740	0.834	3 071 763	8.1
\	\	√	0.876	0.781	0.845	3 005 843	8.1
√	√	\	0.880	0.757	0.843	3 092 883	8.2
√	\	√	0.895	0.743	0.837	3 026 963	8.1
\	√	√	0.883	0.754	0.847	3 071 763	8.1
√	√	√	0.897	0.794	0.854	3 077 043	8.3

注：“√”表示加入该模块；“\”表示没有加入该模块。

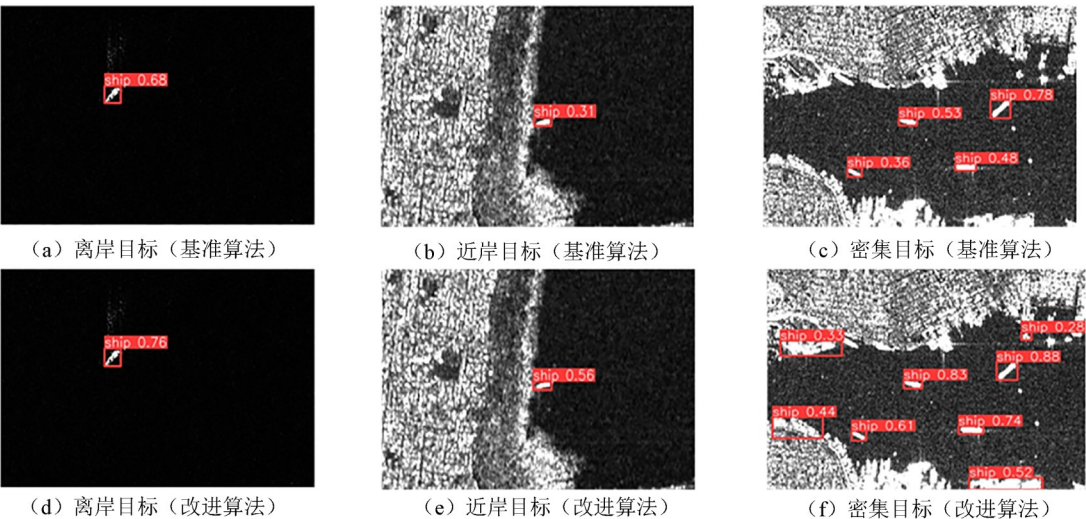


图 4 检测效果对比

Fig.4 Comparison of detection results

表 2 不同算法性能实验结果
Tab. 2 Performance test results of different algorithms

方法	$P/\%$	$R/\%$	$P_{mAP50}/\%$	参数量/个	每秒 10 亿次 浮点运算数/次
SSD	0.860	0.684	0.794	23 746 127	30.4
RT-DETR	0.819	0.769	0.856	31 985 795	108.1
YOLOv5s	0.841	0.780	0.829	70 122 822	15.8
YOLOv6	0.862	0.745	0.811	4 233 843	11.8
YOLOv7tiny	0.850	0.708	0.767	6 014 038	13.2
YOLOv8s	0.899	0.751	0.852	11 125 971	28.4
本文方法	0.897	0.794	0.854	3 077 043	8.3

3 结论

本文基于 YOLOv8 提出了一种融合注意力机制与全局信息的 SAR 舰船目标检测算法。通过将感受野注意力机制与传统卷积相结合,提高了骨干网络的特征获取能力;优化了空间金字塔池化结构,实现局部特征与全局特征的融合,有效提高了模型的感受野和鲁棒性;引入了内部交并比损失函数以提高模型的泛化能力,提升了预测框的准确度。最终在参数量和运算速度几乎不变的情况下提升了对舰船目标的检测精度。

通过对比实验证明改进后的算法整体性能明显提升,平均精度上升 3 个百分点。表明本文所作改进能够有效提高对 SAR 图像舰船目标的检测性能,并且具有较高的实时性,可以实现在小型设备上部署的需求,对于海域舰船检测和弹载目标智能识别算法具有实际意义。

参考文献:

[1] DING J, CHEN B, LIU H W, et al. Convolutional neural network with data augmentation for SAR target recognition[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(3): 364-368.

[2] 马啸, 邵利民, 金鑫, 等. 舰船目标识别技术研究进展[J]. 科技导报, 2019, 37(24): 65-78.

MA X, SHAO L M, JIN X, et al. Advances in ship target recognition technology[J]. Science & Technology Review, 2019, 37(24): 65-78.

[3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings

of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 580-587.

[4] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[5] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 936-944.

[6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer, 2016: 21-37.

[7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 779-788.

[8] 张上, 李梦思, 陈永麟, 等. 基于改进 YOLOv7 的 SAR 图像舰船目标检测算法[J]. 电光与控制, 2024, 31(5): 46-53.

ZHANG S, LI M S, CHEN Y L, et al. An improved YOLOv7 based algorithm for ship target detection in SAR images[J]. Electronics Optics & Control, 2024, 31(5): 46-53.

[9] ZHAO K, LU R T, WANG S Y, et al. ST-YOLOA: A Swin-transformer-based YOLO model with an attention mechanism for SAR ship detection under complex background[J]. Frontiers in Neuro Robotics, 2023, 17: 1170163.

- [10] GUO Y, CHEN S Q, ZHAN R H, et al. LMSD-YOLO: A lightweight YOLO algorithm for multi-scale SAR ship detection[J]. *Remote Sensing*, 2022, 14(19): 4801.
- [11] CHAUDHARI S, MITHAL V, POLATKAN G, et al. An attentive survey of attention models[J]. *ACM Transactions on Intelligent Systems and Technology*, 2021, 12(5): 53.
- [12] ZHANG X, LIU C, YANG D G, et al. RFACnv: Innovating spatial attention and standard convolutional operation [EB/OL]. 2023[2024-01-15]. <https://doi.org/10.48550/arXiv.2304.03198>.
- [13] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. New York, USA: AAAI, 2020, 34(7): 12993-13000.
- [14] ZHANG H, XU C, ZHANG S J. Inner-IoU: More effective intersection over union loss with auxiliary bounding box [EB/OL]. 2023[2024-01-15]. <https://doi.org/10.48550/arXiv.2311.02877>.
- [15] LI J W, QU C W, SHAO J Q. Ship detection in SAR images based on an improved faster R-CNN[C]//*2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA)*. Piscataway, NJ, USA: IEEE, 2017: 1-6.
- [16] ZHAO Y A, LV W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection [EB/OL]. 2024[2024-01-15]. <https://doi.org/10.48550/arXiv.2304.08069>.